

United States Patent Application

for

AUTOMATIC DOCUMENT DETECTION METHOD AND SYSTEM

Inventors:

Qian Lin
Daniel Tretter
Jian Fan

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: **EE094733458US**

Date of Deposit: **October 17, 2001**

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

ERIC HO

(Typed or printed name of person mailing paper or fee)

Eric Ho

(Signature of person mailing paper or fee)

AUTOMATIC DOCUMENT DETECTION METHOD AND SYSTEM

FIELD OF THE INVENTION

The present invention relates generally to digital cameras, and more particularly, to a method and system for automatically determining that a scene is a document and tailoring the image capture and image processing accordingly.

BACKGROUND OF THE INVENTION

Most digital cameras have a single mode of operation, and as such, do not provide any special processing for documents. As can be appreciated, since the same image processing techniques and image capture parameters are uniformly applied to the capture image without regard to the content of the image, documents captured by these digital cameras are of very poor quality and are often not readable.

There are some cameras, such as the RDC-i700 digital camera available from Ricoh Inc. of West Caldwell, New Jersey, that have a document mode. With these cameras, a user can manually select a document mode. Once in document mode, the camera attempts to use camera settings that are more suitable for documents versus a natural scene.

Unfortunately, the user has to switch the digital camera into document mode. While a user is very good at determining whether a scene is a document, the user may forget to switch to normal mode when taking the next picture. As can be appreciated, this requirement for the user to remember to switch between normal mode and document mode can lead to poor image quality for those natural scenes, where the setting is document mode. Consequently, it would be desirable for there to be a mechanism that would automatically detect whether a scene is a natural scene or a document and automatically switch to an appropriate mode with user intervention.

Furthermore, those cameras with a document mode offer only tolerable and primitive image processing that leads to very noisy images. For example, the

Figure 1 consists of 15 bar charts, labeled (a) through (o), each representing a different variable. The y-axis for all charts is 'percentage of total sample' ranging from 0 to 100. The x-axis for all charts has four categories: 'no', 'small', 'big', and 'very big'. The variables are: (a) age, (b) sex, (c) marital status, (d) number of children, (e) number of siblings, (f) number of parents, (g) number of grandparents, (h) number of great-grandparents, (i) number of great-great-grandparents, (j) number of great-great-great-grandparents, (k) number of great-great-great-great-grandparents, (l) number of great-great-great-great-great-grandparents, (m) number of great-great-great-great-great-great-grandparents, (n) number of great-great-great-great-great-great-great-grandparents, and (o) number of great-great-great-great-great-great-great-great-grandparents. The data shows that for most variables, the 'no' category is the most frequent, while for some variables like 'age' and 'sex', the 'small' and 'big' categories are more frequent.

Figure 1 consists of 15 bar charts, labeled (a) through (o), each representing a different variable. The y-axis for all charts is 'percentage of total sample' ranging from 0 to 100. The x-axis for all charts has four categories: 'no', 'small', 'big', and 'very big'. The variables are: (a) age, (b) sex, (c) marital status, (d) number of children, (e) number of siblings, (f) number of parents, (g) number of grandparents, (h) number of great-grandparents, (i) number of great-great-grandparents, (j) number of great-great-great-grandparents, (k) number of great-great-great-great-grandparents, (l) number of great-great-great-great-great-grandparents, (m) number of great-great-great-great-great-great-grandparents, (n) number of great-great-great-great-great-great-great-grandparents, and (o) number of great-great-great-great-great-great-great-great-grandparents. The data shows that for most variables, the 'no' category is the most frequent, while for some variables like 'age' and 'sex', the 'small' and 'big' categories are more frequent.

SUMMARY OF THE INVENTION

According to one embodiment of the present invention, an automatic document detection method is described. First, a preview image of a scene is captured. Next, an automatic determination is made whether the scene is a document. When it is determined that the scene is a document, at least one camera control is programmed with a value that is tailored for document capture. The scene is then captured using the programmed camera controls. Image processing that is tailored for documents is then performed on the captured scene. When it is determined that the scene is not a document, standard camera settings are used for image capture, and standard image processing is performed on the captured image.

10006289-1

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements.

5 FIG. 1 illustrates a digital image capture device in which the automatic document detection mechanism and document image processing mechanism according to one embodiment of the present invention can be utilized.

10 FIG. 2 is a block diagram that illustrates in greater detail the automatic document detection mechanism in accordance with a preferred embodiment of the present invention.

FIG. 3 is a block diagram illustrating in greater detail the document image processing mechanism in accordance with one embodiment of the present invention.

15 FIG. 4 is a flow chart illustrating the processing steps performed by the automatic document detection mechanism of FIG. 2 in accordance with one embodiment of the present invention.

FIG. 5 is a flow chart illustrating the processing steps for automatic document detection in accordance with an alternative embodiment of the present invention.

20 FIG. 6 is a flow chart illustrating the processing steps performed by the document image processing mechanism of FIG. 3 in accordance with one embodiment of the present invention.

25 FIG. 7 is a block diagram illustrating in greater detail the document mode camera control mechanism in accordance with one embodiment of the present invention.

FIG. 8 and FIG. 9 illustrate vertical differences and horizontal differences, respectively, that may be utilized in step 410 of FIG. 4.

DETAILED DESCRIPTION

A method and system for automatically determining that a scene is a document and tailoring the image capture and image processing for documents are described. In the following description, for the purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form in order to avoid unnecessarily obscuring the present invention.

Taking a picture involves letting light fall on film or an image sensor under controlled conditions. This process is often referred to as an exposure. When a photographer presses the shutter button, blades (known as a diaphragm) inside the lens shift to form an opening that is referred to as the aperture. As can be appreciated, the amount of light that exposes a frame depends on the shutter speed and the size of the aperture.

As described previously, the lens has diaphragm blades that open and close to form certain-sized holes (i.e., apertures) that control the amount of light allowed to expose the film or image sensor. The aperture scale, which is found on the lens' aperture ring, is referred to as f-number or f/stops.

In addition to controlling the quantity of light entering the camera, the aperture affects the depth of field, which in turn affects the way that a picture looks. When a subject is in focus, there is a certain area in front of the subject and behind the subject that is also in focus. This range of sharpness is called depth of field.

One or more of these different parameters may be controlled by a document mode camera control mechanism 134 of the present invention as described in greater detail hereinafter with reference to FIG. 7.

Digital Image Capture Device 100

FIG. 1 illustrates a digital image capture device 100 in which the automatic document detection mechanism (ADDM) 110 and document image processing mechanism 120 according to one embodiment of the present invention can be utilized. As used herein, the term "document" can be, but is not limited to, a magazine page, a page in a book, a computer printout, information written on a whiteboard, a slide projected from a projector (e.g., a LCD projector), a presentation displayed by a projector (e.g., an overhead projector). A document can include a mixture of text, graphics, and images.

The digital image capture device 100 includes an automatic document detection mechanism (ADDM) 110 for automatically evaluating whether a scene is a document or a natural scene (i.e., a non-document image). The automatic document detection mechanism 110 is described in greater detail hereinafter with reference to FIGS. 2, 4, and 5.

The digital image capture device 100 includes a document processing block 130 for processing scenes that are determined to be a document by the automatic document detection mechanism 110 and a natural scene processing block 150 for processing scenes that are determined not to be a document by the automatic document detection mechanism 110.

The document processing block 130 includes a document specific camera control unit 134 for providing camera settings to optimize the capture of a document. The document processing block 130 also includes a capture unit 138 for capturing the document. The document processing block 130 also includes a document image processing unit 144 for applying image processing algorithms that are tailored for enhancing document images.

The natural scene processing block 150 includes a natural scene specific camera control unit 154 for providing camera settings to optimize the capture of natural scenes (i.e., non-document images). The natural scene processing block 150

also includes a capture unit 158 for capturing the natural scene. The natural scene processing block 150 also includes a image processing unit 164 for applying image processing algorithms that are tailored for enhancing natural scenes.

It is noted that the capture unit 138 and the capture unit 158 may be implemented with a single image capture unit that has different settings as described in greater detail hereinafter with reference to FIG. 7. Similarly, the document processing block 130 and the natural scene processing block 150 may be implemented by a single image processing unit that executes different image processing programs as described in greater detail hereinafter with reference to FIG.

6. It is further noted that the document specific camera control unit 134 and the natural scene specific camera control unit 154 may be implemented as a single control unit that controls the image capture unit and the image processing unit.

Automatic Document Detection Mechanism 110

FIG. 2 is a block diagram that illustrates in greater detail the automatic document detection mechanism 110 in accordance with a preferred embodiment of the present invention. The automatic document detection mechanism 110 includes an image divider 210 for dividing an image (e.g., the preview image) into a plurality of regions and an edge detector 220 (e.g., a luminance edge detector) for detecting the luminance edges in each region. An edge pixel counter 230 is provided for counting the number of luminance edges in each region.

A region determination unit 240 (also referred to herein as a region counter) determines the number of regions in which the number luminance edges is greater than a predetermined number of edges. When the number of regions exceeds a predetermined number of regions, the classifier 250 classifies the image as a document. Otherwise, when the number 244 of regions does not exceed the predetermined number 248 of regions, the classifier 250 classifies the scene as a

non-document. A Boolean variable or flag (e.g., a document flag 254) may be employed to denote whether an image is classified as a document or non-document

Any document image identification algorithm that is tailored for operation on images captured by a digital camera can be utilized to detect whether the current scene is a document. A preferred document image identification algorithm is now described with reference to FIG. 4. An alternative document image identification algorithm is described with reference to FIG. 5.

FIG. 4 is a flow chart illustrating the processing steps performed by the automatic document detection mechanism of FIG. 2 in accordance with one embodiment of the present invention. In step 410, luminance edges within the image are detected. It is noted that edge detection algorithms that are well known by those of ordinary skill in the art can be utilized in step 410.

An exemplary edge detection scheme is now described. For each pixel location, calculate metric of adjacent differences (D1) based on image luminance Y and then compare D1 with a predetermined threshold value (Th) (e.g., 400). When the metric value (D1) is greater than the predetermined threshold (Th), the pixel is classified as an edge pixel. Otherwise, the pixel is classified as a non-edge pixel.

The metric of adjacent differences (D1) can be expressed as follows:

$$\begin{aligned}
 D1 = & |Y_{22} - Y_{12}| + |Y_{32} - Y_{22}| + |Y_{42} - Y_{32}| + |Y_{52} - Y_{42}| + \\
 & |Y_{23} - Y_{13}| + |Y_{33} - Y_{23}| + |Y_{43} - Y_{33}| + |Y_{53} - Y_{43}| + \\
 & |Y_{24} - Y_{14}| + |Y_{34} - Y_{24}| + |Y_{44} - Y_{34}| + |Y_{54} - Y_{44}| + \\
 & |Y_{22} - Y_{21}| + |Y_{23} - Y_{22}| + |Y_{24} - Y_{23}| + |Y_{25} - Y_{24}| + \\
 & |Y_{32} - Y_{31}| + |Y_{33} - Y_{32}| + |Y_{34} - Y_{33}| + |Y_{35} - Y_{34}| + \\
 & |Y_{42} - Y_{41}| + |Y_{43} - Y_{42}| + |Y_{44} - Y_{43}| + |Y_{45} - Y_{44}|
 \end{aligned}$$

FIG. 8 illustrates vertical differences that may be utilized in step 410 of FIG. 4. FIG. 9 illustrates horizontal differences that may be utilized in step 410 of FIG. 4.

In step 420, the distribution (e.g., the spatial distribution) of the edge locations is evaluated. The evaluation step of 420 can include the following sub-steps. In step 430, the image is divided into regions (e.g., rectangular regions of equal size). In step 440, the number of edge pixels within each region is counted. In step 450, the number (T_w) of regions with edge count that is more than a predetermined edge count is determined. For example, the predetermined edge count may be expressed as a percentage (e.g., 50%) of the total region size.

In decision block 454, a determination is made whether the number of regions (T_w) is greater than a predetermined number of regions. The predetermined number of regions may be expressed as a predetermined percentage (e.g., 60%) of the total number of regions in the image. In step 460, when T_w is larger than a predetermined percentage of the total number of regions, the image is classified as a document type. Otherwise, in step 470 the image is classified as non-document type (e.g., a natural scene).

FIG. 5 is a flow chart illustrating the processing steps for automatic document detection in accordance with an alternative embodiment of the present invention. An alternative manner in which to determine whether a scene is natural or a document is now described. In step 510, every pixel is classified into three classes of pixels, such as a text pixel class, a picture pixel class, and a background pixel class. In step 520, the number of text pixels is counted.

In decision block 530, a determination is made whether the number of text pixels is in a predetermined relationship with a predetermined percentage of the total pixels (e.g., it is determined whether the number of text pixels is larger than a predetermined percentage of the total pixels). It is noted that in step 530 the

predetermined percentage can be derived by empirical tests on different types of documents.

When the number of text pixels is in a predetermined relationship with a predetermined percentage of the total pixels, in step 540, the image is classified as a document. Otherwise, when the number of text pixels is not in a predetermined relationship with a predetermined percentage of the total pixels, in step 550, the image is classified as a non-document.

Document Image Processing Mechanism 120

FIG. 3 is a block diagram illustrating in greater detail the document image processing mechanism 144 in accordance with one embodiment of the present invention. The document image processing mechanism 144 includes an edge pixel detector 310 for detecting edge pixels, a sharpening module 320 for sharpening the edge pixels, and a darkening module 330 for darkening the edge pixels. A luminance correction unit 340 corrects luminance of the image. For example, this may involve estimating an illumination map using the edges detected within regions and then correcting for the varying illumination across the input image.

FIG. 6 is a flow chart illustrating the processing steps performed by the document image processing mechanism of FIG. 3 in accordance with one embodiment of the present invention. The edge pixels of the image are first identified. For example, the edge pixels may be identified by the processing of FIG. 4 or the processing of FIG. 5. In step 610, the edge pixels determined by the processing of FIG. 4 or text pixels determined by the processing of FIG. 5 are sharpened and darkened. This image processing is tailored for documents and makes the text, graphics, and images of a document more readable by sharpening the edges of text and also by darkening the text.

In step 620, luminance correction is performed on the image. Since the lighting for the image capture may be non-uniform, the document specific image

processing unit 144 corrects for non-uniformities in the background. For example, the background pixels that can represent the paper on which the text is printed may be non-white, where in fact the background of the document is supposed to be white. In this case, the document specific image processing corrects these pixels to reflect the background of the document.

Document Mode Camera Control Mechanism 140

In one embodiment, the automatic flash control unit 710 disables the flash to tailor the image capture for documents. It is noted that when a flash is utilized and directly pointed at the document, there is severe glare in the image such that a portion of the image tends to become washed out (e.g., a white area) regardless of the content of the document.

Furthermore, the shutter speed control unit 720 sets the shutter speed at a predetermined shutter speed (e.g., 1/30 second or faster) in order to avoid possible motion blur caused by movement of a user's hand during image capture.

The aperture control unit 730 determines an appropriate aperture setting based on the predetermined shutter speed. For example, the aperture control unit 730 can set an aperture with a maximum opening, corresponding to a minimum f-number. When a maximum aperture is not available for the image capture device, the ISO control unit 740 accommodates the current lighting situation by modifying the ISO film speed (e.g., by increasing the ISO film speed).

FIG. 7 is a block diagram illustrating in greater detail the document mode camera control mechanism in accordance with one embodiment of the present invention. The document mode camera control mechanism includes an automatic flash control unit 710, a shutter speed control unit 720, an aperture control unit 730, an ISO control unit 740, and a user interface unit 750 for generating messages and instructions for the user.

In one embodiment, the digital camera includes an automatic flash, a shutter speed control, an aperture control, and a capture plane. The document to be captured is disposed in a document plane. A first example of the type of settings programmed by the document mode camera control mechanism of FIG. 6 is now described. In this example, a user is instructed to position the digital camera in a first predetermined manner, where the capture plane is approximately parallel to the document plane. Then, the document mode camera control mechanism 134 employs the automatic flash control unit 710 to disable the automatic flash. Alternatively, the document mode camera control mechanism 134 can utilize the user interface unit 750 for generating a message to instruct the user to manually disable the flash.

The shutter speed control unit 720 sets the shutter speed to a predetermined shutter speed (e.g., 1/30 second or faster). Then, the aperture control unit 730 determines an appropriate aperture setting based on the selected shutter speed. When the required aperture is beyond the range of available aperture settings, the ISO control unit 740 can modify the ISO film speed in order to accommodate a wide variety of possible lighting situations.

In a second example, the user is instructed to position the digital camera in a second predetermined manner that reduces reflections from the document. In this example, the capture plane is at an angle with respect to the document plane, and the angle may have a value in a predetermined range of angle values. Preferably, the predetermined range of angle values includes the range from about 22 degrees to about 45 degrees. The document mode camera control mechanism enables the automatic flash and sets a small aperture with a f-number greater than or equal to a predetermined value, for example, f/5.6, in order to secure enough depth of field to avoid out of focus blur. The document mode camera control mechanism then determines a shutter speed based on the aperture setting and sets the shutter speed accordingly.

In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident that various modifications and changes may be made thereto without departing from the broader scope of the invention. The specification and drawings are, accordingly, to be
5 regarded in an illustrative rather than a restrictive sense.

02323660